

ANALISIS SENTIMEN TWITTER TERHADAP KENAIKAN BAHAN BAKAR MINYAK MENGGUNAKAN ALGORITMA NAÏVE BAYES

Muhammad Ghifari^{1*}, Reva Ragam Santika², Sejati Waluyo³

^{1,2,3} Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, DKI Jakarta, Indonesia

Email: ^{1*}1811502945@student.budiluhur.ac.id, ²reva.ragam@budiluhur.ac.id, ⁴sejati.waluyo@budiluhur.ac.id
(* : corresponding author)

Abstrak-Perkembangan zaman yang semakin canggih berefek langsung kepada perilaku masyarakat, di zaman sekarang ini pula penggunaan sosial media semakin menjamur di tengah-tengah aktivitas masyarakat. Khususnya sosial media seperti Twitter yang banyak digemari pada saat ini Twitter menyediakan layanan bagi pengguna untuk menyampaikan pendapat dan mengirimkan pesan secara langsung melalui postingan dan tweet.. Melimpahnya opini masyarakat yang tertuang dalam media sosial twitter berupa sebuah persepsi, baik itu positif atau negatif. Di sisi lain, opini publik juga banyak digunakan sebagai bahan penelitian untuk mencari informasi yang membutuhkan teknik analisis yang tepat, dan informasi yang dihasilkan membantu banyak pihak yang berkepentingan untuk mengambil keputusan. Oleh karena itu, dalam penelitian ini, dengan mengklasifikasikan sentimen menjadi positif dan negatif, kami menggunakan metode algoritma Naive Bayes untuk menganalisis sentimen guna mendeteksi persepsi masyarakat terhadap isu kenaikan BBM di media sosial Twitter. Klasifikasi dilakukan dengan menggunakan metode algoritma naïve bayes sebagai pengujian. Hasil dari penelitian ini adalah dari perbandingan data latih dan uji yaitu 90:10 dan mendapatkan nilai akurasi tertinggi 81,818% dibandingkan dengan rasio lainnya. Perkiraan yang diperoleh dari pengelompokan lebih besar untuk label positif. Selain melakukan klasifikasi untuk menggambarkan tampilan, penelitian ini juga mengevaluasi kinerja model yang dihasilkan.

Kata Kunci: Sentimen Analisis, Ulasan Twitter, Kenaikan BBM, Naïve Bayes.

SENTIMENT ANALYSIS OF TWITTER ON FUEL OIL INCREASE USING NAÏVE BAYES ALGORITHM

Abstract- The development of an increasingly sophisticated era has a direct effect on people's behavior, in this day and age the use of social media is increasingly mushrooming in the midst of community activities. Especially social media such as Twitter which is much in vogue at this time. Twitter provides services for users to express opinions and send messages directly through posts and tweets. On the other hand, public opinion is also widely used as research material to find information that requires appropriate analytical techniques, and the information produced helps many interested parties to make decisions. For example, like what happened recently where the increase in fuel oil (BBM) was widely discussed, causing pros and cons in the midst of society from all fields. Therefore, in this study, by classifying sentiment into positive and negative, we use the Naive Bayes algorithm method to analyze sentiment to detect public perceptions of the fuel price hike issue on Twitter social media. Classification is carried out using the naïve Bayes algorithm method as a test. The results of this study are from a comparison of training and test data, namely 90:10 and obtaining the highest accuracy value of 81.818% compared to other ratios. Estimates obtained from larger groupings for positive labels. In addition to classifying to describe the appearance, this study also evaluates the performance of the resulting model.

Keywords: Sentiment Analysis, Twitter Reviews, Fuel Rises, Naïve Bayes.

1. PENDAHULUAN

Bahan bakar minyak (BBM) di Indonesia merupakan salah satu aspek penting di kehidupan masyarakat, bahkan saat ini bisa terbilang merupakan kebutuhan pokok bagi masyarakat terutama yang memiliki transportasi darat. Pentingnya kondisi keadaan BBM di Indonesia juga faktor yang sangat mempengaruhi banyak hal secara langsung maupun tidak langsung, salah satu yang terpenting seperti yang saat ini terjadi di segi sektor transportasi, mulai dari pendistribusiannya, produksi, bahkan berefek langsung ke sektor lainnya seperti rumah tangga dan yang lainnya. Baru saja di Indonesia terjadi isu tentang kenaikan harga BBM di Indonesia, dan pastinya hal ini banyak menimbulkan pro kontra di tengah masyarakat, tidak hanya kalangan menengah kebawah melainkan masyarakat kalangan atas juga banyak yang berkomentar terkait hal ini.

Kenaikan BBM saat ini sebagian dipicu karena adanya konflik yang terjadi di luar negeri yaitu rusia dan ukraina. Kenaikan harga BBM saat ini tidak hanya berdampak pada masyarakat miskin, tetapi juga merupakan masalah anggaran nasional yang penting yang tidak dapat dipisahkan dari adanya subsidi yang diberikan pemerintah untuk permintaan minyak. Banyak opini masyarakat yang timbul terkait hal ini di berbagai media sosial salah satu yang terbesar yaitu Twitter, masyarakat menyampaikan keluh kesahnya juga banyak memberikan

kritik serta saran untuk pemerintah dalam pengambilan keputusannya terkait hal ini tidak hanya opini negatif tetapi banyak juga opini komentar positif masyarakat yang mendukung demi kebaikan dan manfaat di berbagai hal.

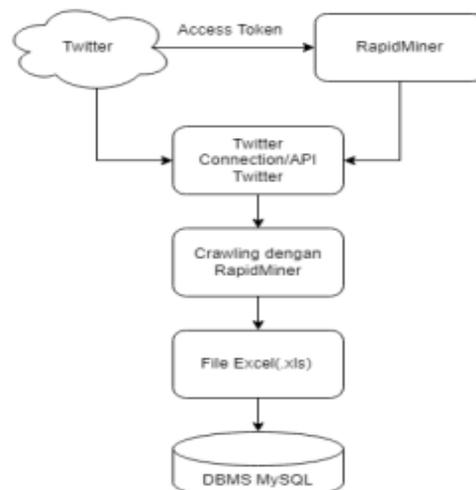
Pembahasan studi ini berfokus pada analisis sentimen komunitas di Twitter terkait kenaikan BBM di Indonesia baru-baru ini. Analisis sentimen adalah ilmu dalam menganalisis penilaian seseorang terhadap suatu hal yang biasanya berupa pendapat, sikap, dan emosi. Di penelitian ini pesan komentar masyarakat yang membahas terkait isu kenaikan harga BBM akan dikumpulkan terlebih dahulu yang berasal dari media sosial twitter, selanjutnya pesan atau komentar tersebut diolah apakah pendapat atau pesan tersebut bersifat positif, negatif, netral. Tujuannya demi menggali opini masyarakat terkait kebijakan kenaikan harga BBM, yang nantinya hasil penelitian menunjukkan persentase dari opini atau sentimen masyarakat umum mengenai hal adanya kenaikan BBM.

2. METODE PENELITIAN

Dalam perancangan dan proses analisis sentimen pada penelitian ini dibagi melalui beberapa tahap. Seperti proses-proses yang diawali dari pengambilan data (crawling), preprocessing data, pelabelan data, pembagian data, perhitungan algoritma naïve bayes serta pengujian data.

2.1 Pengumpulan Data (*Crawling*)

Pada penelitian ini pengambilan data dilakukan dari salah satu platform media sosial yaitu twitter. Proses Crawling ini dilakukan untuk mengumpulkan komentar atau tweet dari twitter terkait isu kenaikan BBM. Proses Crawling ini adalah proses mengumpulkan data pada suatu halaman atau platform tertentu yang tujuannya data tersebut nantinya akan dianalisis dan di proses lebih lanjut. Dalam penelitian ini pengambilan data dilakukan menggunakan RapidMiner yang berbasis GUI (graphical user interface) yang terhubung menggunakan API Twitter.



Gambar 1. Tahapan Pengumpulan Data

2.2 Preprocessing

Preprocessing adalah proses untuk membersihkan data dari objek yang tidak diperlukan pada suatu dokumen atau kalimat untuk bertujuan mempermudah proses pengolahan data. Selain itu proses ini juga berguna untuk menghapus data yang kurang lengkap, kejanggalan data, dan data yang tidak utuh atau stabil. Data yang berupa teks juga memerlukan proses konversi untuk mengubah teks.

2.3 Pelabelan Data

Pelabelan data adalah proses mengklasifikasikan kalimat berdasarkan fitur yang dikandungnya. Ada dua kelas sentimen yang dievaluasi dalam proses ini, sentimen positif dan negatif. Hasil skoring memberikan data evaluasi sentimen positif dan negatif.

2.4 Pembagian Data

Pada tahapan pembagian data, tweet yang telah dilabeli akan dibagi dua buah segmen yaitu data latih dan data uji. Pembagian data dilakukan dengan membagi dataset menjadi 90% data latih dan 10% data uji. Data latih adalah kumpulan data yang berasal dari proses atau tahap preprocessing. Sedangkan Data uji adalah data yang disiapkan untuk mengukur tingkat keakuratan algoritma klasifikasi berdasarkan model yang dilatih.

2.5 Naïve Bayes

Pengacuan Naïve bayes adalah algoritma yang berkemampuan menerima input dalam bentuk apapun dan mempunyai kecepatan dalam mengoperasikan suatu data. Berdasarkan kompleksitasnya, Naive Bayes merupakan salah satu algoritma yang paling sederhana untuk menerapkan aturan Bayesian dengan beberapa keunggulan yaitu sangat efisien, membutuhkan data eksperimen yang sedikit, mudah diimplementasikan dan memiliki akurasi yang relatif tinggi. Algoritma Naive Bayes berdasarkan teorema Bayes bahwa semua operasi membuat kontribusi yang sama besar atau independen untuk pemilihan kelas tertentu. Naïve Bayes dapat dilatih secara efektif dalam pembelajaran yang diawasi, proses klasifikasi Naïve Bayes telah ramai digunakan dalam riset yang bertujuan untuk mengelompokan data dalam jumlah yang besar. Manfaat dari pengklasifikasi adalah hanya membutuhkan beberapa kecil data pelatihan untuk melihat para-parameter (sarana dan varians dari variable) yang diperlukan untuk pengelompokan.

$$p(C_k | \mathbf{x}) = \frac{p(C_k) p(\mathbf{x} | C_k)}{p(\mathbf{x})} \quad (1)$$

Keterangan :

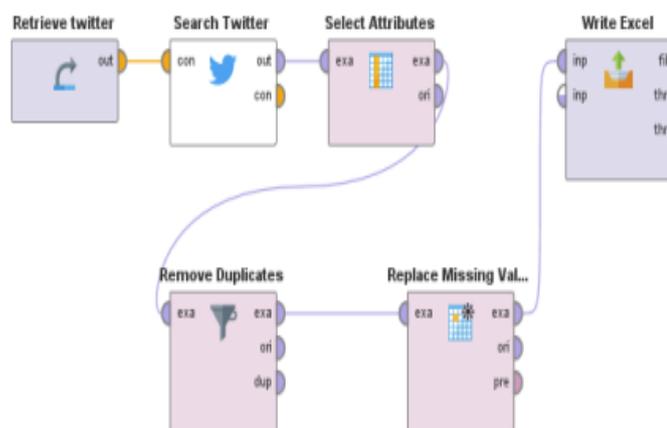
- $p(C_k | \mathbf{x})$ = Probabilitas hipotesis C_k jika diberikan fakta atau *record* X (*Posterior probability*).
- $p(\mathbf{x} | C_k)$ = Mencari nilai parameter yang memberi kemungkinan yang paling besar (*likelihood*).
- $p(C_k)$ = *Prior probability* dari X (*Prior probability*).
- $p(\mathbf{x})$ = Jumlah *probability tuple* yang muncul.

3. HASIL DAN PEMBAHASAN

Pada bagian ini berfokus pada tahap yang dilakukan, beberapa tahap yaitu crawling data (pengumpulan data), *preprocessing*, pelabelan data, pembagian data, Pengujian Data serta perhitungan nilai akurasi. Tahapan yang akan dilalui dalam penelitian ini.

3.1 Crawling Dataset

Crawling data diambil dari platform media sosial twitter dengan memanfaatkan API Twitter. Data yang digunakan 500 data dan kata kunci pencarian yaitu kenaikan bbm dan hanya mengambil data yang berbahasa Indonesia dengan tools rapidminer. Pada proses tersebut dilakukan didalam rapidminer berikut adalah langkah crawling di rapidminer.



Gambar 2. Crawling Dataset

- a. Operator *Retrieve* twitter untuk menyambungkan RapidMiner dengan Twitter.
- b. Operator *Search* Twitter untuk mencari data (tweet) dari kata kunci “kenaikan bbm”.
- c. Operator *Select Attributes* untuk memilih atribut yang ingin diperlukan. Dalam penelitian ini diperlukan atribut text (tweet).
- d. Operator *Remove Duplicate* untuk menghapus Tweet yang sama (spam).
- e. Operator *Replace Missing Value* untuk menghapus Tweet yang tidak lengkap.
- f. Operator *Write Excel* untuk menulis (*export*) data ke file excel (.xlsx).

3.2 Tahap Preprocessing

Text preprocessing adalah proses pemilihan dan penataan lebih lanjut data teks melalui serangkaian tahapan seperti case folding, tokenization, filtering, dan stemming. Namun, dalam praktiknya, tidak ada aturan yang jelas untuk setiap tahapan text preprocessing. Itu semua tergantung pada sifat dan keadaan data yang kita miliki. Proses Preprocessing antara lain.

- a. **Case folding** dilakukan penyetaraan teks menjadi huruf kecil secara keseluruhan, misalnya, “BBM” akan diubah menjadi “bbm”, “Kenaikan” akan diubah menjadi “kenaikan” dan seterusnya.

Tabel 1. Tahap Case Folding

Sebelum	Sesudah
Kebijakan kenaikan harga BBM jaga keseimbangan ekonomi Indonesia	kebijakan kenaikan harga bbm jaga keseimbangan ekonomi indonesia
KITA MENOLAK BBM NAIK	kita menolak bbm naik

- b. **Cleansing** yaitu proses membersihkan data dari noise seperti hashtag, username, url, tanda baca, dan angka. Terdapat beberapa proses antara lain.

1. **Remove Punctuation** berfungsi untuk menghilangkan tanda baca teks.

Tabel 2. Remove Punctuation

Sebelum	Sesudah
Pksejahtera menolak kenaikan bbm tapi begitu ada blt bbm antri pertama????	Pksejahtera menolak kenaikan bbm tapi begitu ada blt bbm antri pertama
Cnnindonesia demo rakyat + mahasiswa berhari2 menentang	Cnnindonesia demo rakyat mahasiswa berhari2 menentang

- c. **Remove Username** berfungsi menghilangkan username biasanya diawali dengan simbol ‘@’. Dikarena pada hal ini dianggap tidak penting.

Tabel 3. Remove Username

Sebelum	Sesudah
@pksejahtera menolak kenaikan bbm tp begitu ada blt bbm antri pertama???????	menolak kenaikan bbm tp begitu ada blt bbm antri pertama???????
@Cnnindonesia itu demo rakyat + mahasiswa berhari2 menentang kenaikan bbm aja presiden ga takut koq??	tu demo rakyat + mahasiswa berhari2 menentang kenaikan bbm aja presiden ga takut koq??

- d. **Remove Hashtag** berfungsi menghapus judul topik yang dibahas, dan kemampuan untuk mengelompokkan percakapan, biasanya dengan tanda ‘#’.

Tabel 4. Remove Hashtag

Sebelum	Sesudah
Jalan merdeka barat ditutup karena demo kenaikan bbm #tolakbbmnaik #info	Jalan merdeka barat ditutup karena demo kenaikan bbm
Rakyat harus cerdas menerima kenaikan bbm non subsidi #bbmdiindonesia	Rakyat harus cerdas menerima kenaikan bbm non subsidi

- e. **Clean Number** berfungsi untuk menghapus angka yang berada di depan atau awal kalimat maupun yang ditengah kalimat.

Tabel 5. Clean Number

Sebelum	Sesudah
Cnnindonesia itu demo rakyat + mahasiswa berhari2 menentang kenaikan bbm aja presiden ga takut koq??	Cnnindonesia itu demo rakyat + mahasiswa berhari menentang kenaikan bbm aja presiden ga takut koq??
pada hari kamis tanggal 20 oktober 2022 sekira pukul 09.10 wib personil polsektanjung beringin melaksanakan kegiatan atas pengumuman pemerintah tentang kenaikan harga bbm	pada hari kamis tanggal oktober sekira pukul wib personil polsek tanjung beringin melaksanakan kegiatan atas pengumuman pemerintah tentang kenaikan harga bbm

- f. *Removal* URL berfungsi menghapus tautan, di dalam tweet sering ditampilkan dalam tweet dan URL perlu dihapus, yang membuat data tidak valid. URL biasanya digunakan untuk mempromosikan sesuatu sehingga pengguna lain dapat mengakses URL tersebut.

Tabel 6. Removal URL

Sebelum	Sesudah
Inflasi di Pekanbaru masih tinggi karena dipicu akibat kenaikan harga bbm https://t.co/5KwhQmKW0O	Cnnindonesia itu demo rakyat + mahasiswa berhari menentang kenaikan bbm aja presiden ga takut koq??

- g. *Removal* RT (*Retweet*) salah satu fitur twitter yang membantu pengguna atau orang lain dengan cepat berbagi *tweet* dengan semua pengikutnya.

Tabel 7. Removal RT

Sebelum	Sesudah
@askrlfess masalah lesti ini kaya pengalihan isu ferdi sambo sma kenaikan bbm	masalah lesti ini kaya pengalihan isu ferdi sambo sma kenaikan bbm

- h. *Tokenizing*

Tokenizing adalah proses pemecahan teks menjadi bentuk token (bisa berupa huruf, kata, atau kalimat) sebelum dianalisis lebih lanjut. Entitas yang dapat dirujuk sebagai token, seperti kata, angka, simbol, dan tanda baca. Perhatikan contoh berikut ini “Jalan Merdeka Barat ditutup karena ada demo BBM” Setelah proses tokenizing pada kalimat di atas, akan diperoleh delapan kata yaitu: “jalan”, “merdeka”, “barat”, “ditutup”, “karena”, “ada”, “demo”, “bbm”.

3.3 Pelabelan Data

Pelabelan merupakan tahapan yang hanya dapat dilakukan setelah tersedianya satu atau lebih data clean text pada basis data (database) hasil dari tahapan preprocessing. Tahap pelabelan manual ini dilakukan oleh peneliti sebagai validasinya untuk memberikan label ke seluruh data yang berjumlah 210 tweet. Setelah tahap ini data akan mempunyai label negatif dan positif dengan label negatif berjumlah 82 tweet dan positif berjumlah 128 tweet. Tweet yang sudah didapatkan pada tahap crawling akan diberikan kelas positif atau negatif. Kelas positif salah satu bentuk kalimat pernyataan yang memberikan informasi tentang suatu kejadian yang bernilai benar. Sedangkan kelas negatif dimaksudkan untuk teks tweet kontra dengan kelas negatif.

Tabel 8. Pelabelan Data

Text	Label
nih dengerin fakta mutu pertalite tidak berubah setelah kenaikan harga bbm spesifikasi pertalite tetap	Positif
ciamis siapkan dana miliar untuk warga tidak bagian bantu sosial dan dampak naikan bbm belum saja pulih dari pandemi covid 19 sudah dihajar kenaikan harga bbm dan dampak terusnya	Positif Negatif

Cleansing dilakukan penyaringan dan pembuangan teks. Proses cleansing terdiri atas beberapa tahapan antara lain: menghapus URL, mention (@mention), hashtag (#hashtag), angka atau selain huruf.

Tokenizing adalah proses pemotongan kalimat yang berasal dari tahapan sebelumnya berdasarkan tiap kata penyusunnya. Perhatikan contoh berikut ini. “saya menolak adanya kenaikan harga.” Setelah proses Tokenizing pada kalimat di atas, akan diperoleh lima kata yaitu: “saya”, “menolak”, “adanya”, “kenaikan”, “harga”.

3.4 Pembagian Data

Pada tahapan pembagian data, tweet yang telah berlabel akan dibagi menjadi data uji dan data latih. Pada penelitian ini proses pembagian data dilakukan dengan membagi dataset menjadi 90% data latih dan 10% menjadi data uji.

Tabel 9. Pembagian Data

Jenis Data	Jumlah
Data Latih	189
Data Uji	21
Data Set	210

3.5 Pengujian & Nilai Akurasi

Hasil pengujian pada metode algoritma Naïve bayes dapat dilihat pada Gambar 3.

Pemisahan Data Training & Testing

response	label
--Data Training--	
serendah itu ya oposisi gak ada isu lain apa ada kenaikan bbm tarif listrik dua puluh polisi narkoba ini malah bahas selebar kertas	negatif
yang mengadu ke balai kota dibatasin tidak boleh mengadu soal kenaikan harga sembako bbm dan mahal nya kebutuhan hidup	negatif
gempa bumi banjir bandang tanah longsor kenaikan harga bbm semua gara gara vaksin	negatif
saham emiten consumer goods yang berkaitan dengan industri ritel juga naik jadi narasi kamu tentang dampak kenaikan bbm terhadap ritel salah total	negatif
pastikan ketersediaan bbm cukup babinsa pengamanan dan pengawasan spbu babinsa koramil satu laratama serda abraham tunliu melaksanakan pengamanan dan pengawasan wilayah serta monitoring aktifitas spbu paska terjadinya kenaikan bbm	positif
kritik tiga tahun pemerintah jokowi maruf juru bicara kholid muhammad sampai kritik untuk pemerintah	negatif
mau tau nih alasan pemerintah nurunin harga bbm pertalite sama pertamax dalam kurun waktu berdekatan padahal kemaren memaksa kenaikan karena jebol anggaran	positif
kota pintar penduduknya harus siapin pintar dulu pak anak anak indonesia harus terjamin asupan proteinya kurangi beban pengeluaran rakyat yang kurang agar tidak kekurangan nutrisi karena imbas kenaikan bbm listrik dll karena secara tidak langsung menggerus jatah protein	negatif

Gambar 3. Contoh Data latih

--Data Testing--	
bank indonesia meyakini inflasi pada oktober dua ribu dua dua akan mulai melandai seiring dengan redanya efek rambatan kenaikan harga bbm serta belum kuatnya permintaan	positif
lawan kenaikan bbm dengan mogok umum	negatif
belum saja pulih dari pandemi covid 19 sudah dihajar kenaikan harga bbm dan dampak terusnya	negatif
petugas spbu mogok tuntutan kenaikan upah	negatif
polres pangkalpinang salur bantuan sosial dalam rangka polri peduli kepada masyarakat yang terdampak kenaikan harga bahan bakar minyak bbm aksi peduli itu dipimpin langsung kapolres pangkalpinang	positif
sudah mulai rasakan kenaikan bbm emak emak mengeluh sembako mahal	positif

Gambar 4. Contoh Data Uji

Proses Testing

response	label	Hasil Testing
bank indonesia meyakini inflasi pada oktober dua ribu dua dua akan mulai melandai seiring dengan redanya efek rambatan kenaikan harga bbm serta belum kuatnya permintaan	positif	positif
lawan kenaikan bbm dengan mogok umum	negatif	positif
belum saja pulih dari pandemi covid 19 sudah dihajar kenaikan harga bbm dan dampak terusnya	negatif	negatif
petugas spbu mogok tuntutan kenaikan upah	negatif	positif
polres pangkalpinang salur bantuan sosial dalam rangka polri peduli kepada masyarakat yang terdampak kenaikan harga bahan bakar minyak bbm aksi peduli itu dipimpin langsung kapolres pangkalpinang	positif	positif
sudah mulai rasakan kenaikan bbm emak emak mengeluh sembako mahal	positif	negatif
polsek pamulang melaksanakan pemantauan spbu pasca kenaikan harga bbm di spbu pertamina wilkum pamulang situasi aman dan terkendali	positif	positif
kenaikan harga bbm bikin inflasi indonesia naik lima puluh sembilan persen	negatif	negatif
polsek batu engau menggelar pembagian sembako sebagai bentuk kepedulian dan perhatian polri kepada masyarakat yang terdampak akibat kenaikan bbm bahan bakar minyak)	positif	positif

Gambar 5. Contoh Proses Testing

Setelah data di testing maka nilai sentimen nya ada yang berubah ada juga yang tidak, untuk mengetahui nilai akurasi dari sebuah algoritma naïve bayes menggunakan rumus dibawah. Klasifikasi menggunakan algoritma Naive Bayes diterapkan pada 210 data latih. Prosedur ini berpotensi menghasilkan model yang akan diterapkan pada data uji. Hasil dari proses klasifikasi pada data latih kemudian diaplikasikan pada data uji untuk dijadikan acuan evaluasi klasifikasi menggunakan algoritma Naive Bayes dalam analisis sentimen terhadap kenaikan BBM di twitter.

		<i>Predicted</i>	
		<i>Positive</i>	<i>Negative</i>
<i>Actual</i>	<i>Positive</i>	10	2
	<i>Negative</i>	2	8

Gambar 6. Confusion Matrix

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (2)$$

$$= \frac{10+8}{10+2+8+2} = \frac{18}{22} = 0,818 \times 100\% = 81,8\%$$

Keterangan :

- TP = Jumlah positif yang diklasifikasikan positif
- TN = Jumlah negatif yang diklasifikasikan negatif
- FP = Jumlah negatif yang diklasifikasikan positif
- FN = Jumlah positif yang diklasifikasikan negatif

Pengujian tersebut dilakukan dalam setiap pengembangan sistem untuk mengevaluasi, menganalisis dan mengetahui tingkat akurasi atau kesamaan hasil yang telah dicapai oleh sistem yang telah dirancang. Pada penelitian ini, dilakukan pengujian dari sisi akurasi pada implementasi algoritma Naïve Bayes Classifier (NBC) dalam memprediksikan label untuk data uji. Berikut beberapa pengujian yang dilakukan dengan menggunakan perbandingan 9:1.

Tabel 10. Pengujian Nilai Akurasi

Pengujian	Nilai Akurasi
1	81,81 %
2	77,27 %
3	68,18 %
4	81,81 %
5	77,27 %

4. KESIMPULAN

Berdasarkan hasil pengujian dan evaluasi dari aplikasi yang dibuat menggunakan Dataset dan algoritma yang diusulkan, maka dapat disimpulkan bahwa berdasarkan 210 tweet, sentimen masyarakat Indonesia terhadap isu kenaikan BBM (Bahan Bakar Minyak) dari 5 kali pengujian nilai tertinggi mendapatkan nilai akurasi sebesar 81,81%, Sedangkan nilai terendah mendapatkan nilai akurasi sebesar 68,18%. Tahapan utama yang terdapat dalam penelitian ini antara lain: crawling, preprocessing, pelabelan, pembagian data, dan klasifikasi Naïve bayes. Tahap preprocessing yang baik menjadi penentu dalam terbentuknya hasil yang optimal untuk tahap selanjutnya.

DAFTAR PUSTAKA

- [1] Nooraeni, R., Safiruddin, A. B., Afifah, A. F., Agung, K. D., & Rosyad, N. N. (2020). Analisis Sentimen Publik terhadap Sistem Zonasi Sekolah Menggunakan Data Twitter dengan Metode Naïve Bayes Classification. *Faktor Exacta*, 12(4), 315. <https://doi.org/10.30998/faktorexacta.v12i4.5205>
- [2] Setiawan, H., Utami, E., & Sudarmawan, S. (2021). Analisis Sentimen Twitter Kuliah Online Pasca Covid-19 Menggunakan Algoritma Support Vector Machine dan Naive Bayes. *Jurnal Komtika (Komputasi Dan Informatika)*, 5(1), 43–51. <https://doi.org/10.31603/komtika.v5i1.5189>
- [3] Aditya, D., Mubarak, A., Kom, M., & Susanti, S. (2019). Analisis Sentimen Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes Classifier (Studi Kasus: Komentar Publik Kepada Tri Indonesia). *JURNAL INFORMATIKA*, 6(1), 1–8. <http://ejournal.bsi.ac.id/ejurnal/index.php/ji/article/view/>
- [4] Tanggraeni, A. I., & Sitokdana, M. N. N. (2022). Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma Naïve Bayes. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(2), 785–795. <https://doi.org/10.35957/jatisi.v9i2.1835>
- [5] Saurina, N. (2022). Analisis Sentimen Ulasan Pelanggan Batik Ecoprint Menggunakan Naïve Bayes Dan Knn Classifier. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(2), 1532–1542. <https://doi.org/10.35957/jatisi.v9i2.1483>
- [6] Noor Fahmi, R., Aji Primajaya, dan, Informatika Universitas Singaperbangsa Karawang Jl HSRonggo Waluyo, T., & Timur, T. (2021). Analisis Sentimen Pengguna Twitter Terhadap Kasus Penembakan Laskar FPI Oleh POLRI Dengan Metode Naïve Bayes Classifier. In *Jurnal Informatika dan Komputer* (Vol. 5, Issue 2). <http://dev.twitter.com>.
- [7] Krisdiyanto, T. (2021). Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naïve Bayes Clasifiers. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi*, 7(1), 32. <https://doi.org/10.24014/coreit.v7i1.12945>
- [8] Mutiara, T. A., Alkhalifi, Y., Mayangky, N. A., & Gata, W. (2020). Analisis Sentimen Opini Publik Mengenai Larangan Mudik pada Twitter Menggunakan Naïve Bayes. In *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi* (Vol. 6, Issue 2, p. 85). <https://doi.org/10.24014/coreit.v6i2.9727>
- [9] Sentimen, A., & Algoritma, P. E.-M. P. T. M. (2020). Analisa Sentimen Pengguna E-Money Pada Twitter Menggunakan Algoritma C4.5 dan Naïve Bayes. In *Jurnal Informatika dan Sistem Informasi (JIFoSI)* (Vol. 1, Issue 3).
- [10] Fitriani, E. (2020). Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan. *Sistemasi*, 9(1), 103. <https://doi.org/10.32520/stmsi.v9i1.596>